2019

Reintegrating Biology Through the Nexus of Information and Energy

Kim L. Hoke

Sara L. Zimmer

Mary Jo Ondrechen

Amedee des Georges

Adam B. Roddy

See next page for additional authors

Authors

Kim L. Hoke, Sara L. Zimmer, Mary Jo Ondrechen, Amedee des Georges, Adam B. Roddy, Nicole R. Buan, and Craig E. Williamson

Reintegrating Biology Through the Nexus of Information and Energy

"Open mind for a different view. And nothing else matters"
-Metallica

Kim L. Hoke
Department of Biology, Colorado State University, Fort Collins, CO 80523-1878
Sara L. Zimmer*
Department of Biomedical Sciences, University of Minnesota Medical School, Duluth campus, Duluth, MN 55812
Mary Jo Ondrechen*
Department of Chemistry & Chemical Biology, Northeastern University, Boston, MA, 02115
Amedee des Georges*
Department of Chemistry and Biochemistry, City College of New York, New York, NY 10031
Adam B. Roddy*
Yale School Forestry & Environmental Studies, New Haven, CT 06511
Nicole R. Buan*
Department of Biochemistry, University of Nebraska-Lincoln, Lincoln, NE 68588-0662
Craig E. Williamson
Department of Biology, Miami University, Oxford, OH 45056
*Author order to be randomized

Summary:
Energy and information are two fundamental properties of biological systems at all levels of organization that interact to influence biological responses and their outcomes.

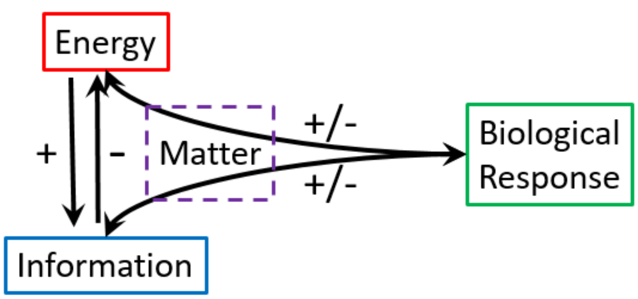

Figure 1: Energy and information are two fundamental properties of biological systems that interact with each other, often, but not always, through matter, to influence biological responses at all levels of organization (e.g. cell, organism, ecosystem).

Full Vision:

The compelling need to reintegrate biology

Recent rapid advances in biology have led to diversification and sub-specialization of many fields, as well as a corresponding explosion of new findings. Advances in tools ranging from genomic techniques and high-resolution imaging to automated ecosystem- and biosphere-level sensors, along with correspondingly advanced analytics, have led to critical new insights that are transforming our understanding of biological systems. One of the consequences of these

Hoke, K.L., Zimmer, S.L., Ondrechen, M.J., desGeorges, A., Roddy, A.B., Buan, N.R., Williamson, C.E. 2019. Reintegrating Biology Through the Nexus of Information and Energy. NSF Reintegrating Biology Jumpstart. https://reintegratingbiology.org/vision-papers/
DOI: 10.1042/ETLS20180031

rapid advances has been accelerated splintering of biology into sub-disciplines with highly focused questions, vocabulary, and techniques. This splintering creates barriers to synergy across fields of biology and hinders the insights that could be gained at their interfaces, creating a compelling need to reintegrate biology. But how can this reintegration be accomplished? A common framework for effective communication across scales of understanding for biological systems is an essential aspect of “reintegration”. Here we propose that **application of fundamental physical and information theory principles** that cross all levels of biological organization would empower the reintegration of biology.

How do energy and information interact and result in outcomes in biological systems?

Information and energy are two fundamental properties of all levels of biological organization. The practical definition of information for the biological sciences that we use here is anything that reduces uncertainty for a biological entity (e.g. DNA, RNA, communication signals between cells or organisms, or different wavelengths of sunlight that stimulate a physiological or behavioral response). We use the term energy to refer to a property of physical or chemical resources that enables work to support life functions such as moving, reproducing, and growing. Fundamentally, life is energy and information organized into material forms. The relationship between energy, information, and matter is therefore key to our understanding of biology.

Understanding the relationships between energy, information, and matter depend on quantifying them in biological systems. In principle, we might quantify **information** at all levels of biological organization. The field of information theory (first developed by Claude Shannon) links information to entropy and, therefore, provides well-defined methods for quantifying information. Entropy is the amount of uncertainty, and information is its reduction. This information theoretic approach has been rigorously applied to ecological systems to quantify species diversity (e.g. Shannon’s Diversity Index and related variants). More recently, macroecologists have developed Maximum Entropy as a method to predict, for example, species-area relationships and species distributions. Information theory is also a central tenet both of sensory neuroscience and of signal detection theory as applied in the fields of psychology and animal behavior and has recently been applied to biochemical and social networks. Extending information theory approaches more broadly could enable more accurate quantification of information in biological systems.

Energy may be quantified at all levels of biological organization as well. For example, at the molecular scale, energy is stored in bonds. Breaking bonds such as those in ATP and carbohydrates requires or releases a defined amount of energy. At the cellular, tissue, or organismal levels, energy is estimated by measuring respiration rate and the change of enthalpy (approximately equivalent to temperature). At the ecosystem scale, energy is quantified by using energy balance equations to estimate radiation inputs and outputs and storage of energy in chemical bonds and biomass (e.g. photosynthetic carbon fixation) and its mobilization and transformation as it moves through an ecosystem. These methods differ in how directly they relate to the actual energy content of a system.

A major challenge is finding common metrics that enable us to link energy or information across levels of biological organization. Energy metrics such as ecosystem biomass, respiration rate, and energy in chemical bonds are challenging to relate quantitatively given the need to account for all chemical bonds in the higher order system. Regarding information, nucleotide order carrying a triplet code has long been recognized as information in DNA, and we have become increasingly aware that there is also information inherent in DNA chromatin state, nucleotide modifications, and protein binding motifs. The information content of biochemical networks, reflected in features like molecular complexity and post-translational modifications to proteins, is not equivalent to the information in the DNA, and the dynamics of these informational sources also differ dramatically. In addition, signals at any level of biological organization might confer minimal benefits when a redundant signal is available. Thus, quantifying energy and information in a biological system is complicated by the context dependence of the expression of information and the propagation of information and information across levels of biological organization.

Theoretically, however, there are ways to address the quantification of information and its interface with energy, and this is where fundamental physical principles come in.

The entropy in a system, S, is a function of the number of states (i) of the system:

$$S = -\sum_i P_i \cdot \log P_i$$

where P_i denotes the probability that the system is in state i.

Because having multiple possible states increases the uncertainty of any one state being occupied, entropy and information (I) are negatively related:

$I = 1 - S$.

Energy and information are inherently physical principles, as increasing information (and reducing entropy) costs energy:

$G = H - TS$

where G is the Gibbs free energy, H is the enthalpy and T is the Kelvin temperature.

The simplicity of equations linking energy and information show the promise of this approach; however, information and energy share a complex relationship, often, but not always involving matter in biological systems. Examples of these complex relationships abound. For example, inherent in an ATP molecule is a specific amount of energy and a measurable amount of information it can convey in cell signalling. Expenditure of either the information or energy is context-dependent. The spectral composition of sunlight on earth is another example, as photosynthetically active radiation (PAR) enables photosynthesis and vision, while ultraviolet radiation (UVR) that is also part of sunlight can cause DNA-damage and organic matter photodegradation. Moreover, UVR is information-rich, providing foraging, mate selection, orientation, migration, phototaxis, and phototropism signals. In summary, energy and information are broadly interrelated yet context-dependent.

Fundamentally, insights into the net effects of energy of various sources on biological systems requires an understanding across multiple levels of biological organization. For example, how do DNA repair processes at the molecular level, and the production of photoprotective compounds or behavioral avoidance of damaging UVR at the physiological level, alter the relative abundance of different species and thus community structure, ecosystem structure and function, and ultimately the integrity of the biosphere? The diverse responses to information by units at one scale can result in an emergent property at a higher hierarchical level, where the response appears to be more than the sum of the individual parts.

Specific benefits to this vision

We envision several benefits of focusing on physical concepts that underlie all biological inquiry. We have described above how we might extract from fundamental physical principles a common framework that provides metrics applicable across levels of biological organization. Currently, a lack of a common framework is preventing clear communication and collaboration among biologists investigating structure, function, or processes at different levels of organization from molecules to the biosphere. Even rudimentary efforts by biologists to identify and quantify the relevant sources of information and energy in the systems that they study will bring us closer to a common, though currently elusive, currency. With this currency in hand, the barriers to interdisciplinary biology research would become reduced and collaborative work would become efficient enough to be attractive and feasible.

Additionally, we believe that such an integrative framework will provide the most direct route to increase our predictive power in biological systems across scales. **The ultimate predictive biological algorithms would ideally account for forces at scales as diverse as atomic to biosphere, therefore back-transformation of biological concepts to their underlying physical and informational properties would result in more accurate predictions in biology.** The greater predictive power from integrating energy mechanisms, information processing, and spatial-kinetic relationships informed by physical principles could have widespread applications, for example breaking the biological “design ceiling” that currently limits our ability to develop the next generation of synthetic organisms that will help address accelerating threats posed by climate change.

Successful establishment of our vision will also result in benefits within a single biological discipline. For example, while universal scaling rules have long been sought to explain biological systems, there remains substantial variation around these scaling relationships, suggesting that other undescribed factors may be needed to explain the system: information could resolve this unknown variation. Similarly, biological systems can function in ways that may not seem logical to the observer (e.g. energy and matter and flow between them are imbalanced), but this apparent paradox may exist precisely because information content and storage has not been considered. These seeming “imbalances” would likely reveal failures by investigators to identify and characterize the information and energetic considerations inherent in the system. Successfully developing methods to quantify information and energy and the relationships between them across scales would help to resolve both known outstanding questions and observed but unquestioned phenomena.

Examples of critical questions that can be addressed include:

- Do RNA and proteins buffer information? What is the energy cost/benefit of transitioning from an RNA world to a DNA world?
- What are the energetic costs and benefits of intra- and intercellular signalling?

- How does information influence organismal structure and function? What epigenetic mechanisms accommodate information input in developmental processes?
- Do ecosystems have an emergent super-organismal ability to capture and respond to information, or is information processed by individual organismal outputs that collectively result in an ecosystem output?
- While energy flow between ecosystems is widely recognized, do ecosystems communicate with each other through information flow as well?
- Do energy and information scale similarly across levels of biological organization (scale-free information-energy relationships), or do levels of biological organization each possess a unique relationship between energy and information (scale-dependent).

Carrying the vision forward: next steps

As with all theoretical and overarching concepts, the immediate challenge is to determine how to implement it. One approach is to immediately implement a plan to **connect (or re-connect) the biological community to physical and information theory principles**. The first concrete step would be to improve the training and engagement of biologists into the fundamental principles of physics and information theory, so that biologists in every field could apply those principles and measure the quantities relevant to describing biological processes in fundamental physics and information theory terms. This would greatly facilitate our understanding of information and energy transfer across scales. While many undergraduate biology curricula include physics and mathematics coursework, rarely, if ever, are the required courses focused on biological applications of information and energetic approaches. The lack of integration in curricula perpetuates the lack of integration of these fields. We also propose a coordinated effort to disseminate examples quantifying energy and information that could be readily adapted by biology professors into their undergraduate courses (e.g. integrating chemistry, physics, and mathematics into case studies for introductory biology, cell biology, and ecology).

Another simple but concrete resource could be a primer of relevant physics for biologists at any career stage. As our efforts to reach across our own disciplines have taught us, such a written resource would be most accessible if it was informed by biophysicists, physicists, and informaticists yet written by biologists who adapt the content for each intended audience using approachable examples. Other possibilities include regular webinars and collaborative workshops to bring together researchers at the graduate, postdoctoral, and early career stages to foster transdisciplinary research and training.